

News in focus

less likely to act. “The ventral pallidum could be the centre for motivation deficit or apathy in depression,” says Amemori.

If confirmed in humans, the findings could shift how clinicians approach one of depression’s most debilitating symptoms. Current treatments often aim to restore enjoyment or reduce anxiety, yet many individuals continue to struggle to start simple tasks. By pinpointing a circuit that selectively dampens motivation in the face of discomfort, the study opens the door to therapies that could focus on lowering that barrier.

Vikram Chib, a biomedical engineer at Johns Hopkins University in Baltimore, Maryland, praises the study’s precision. “Focally disrupting one part of the network and showing a causal effect is important,” he says. “It helps

us understand how this network can go awry.”

Chiu thinks that the study could have implications for psychotherapy, including tailoring cognitive behavioural therapy. Directly modulating this circuit through deep-brain stimulation or non-invasive ultrasound could one day offer another approach, says Amemori.

But he also cautions that the system exists for a reason. Any easing of the motivational brake will require careful balance. “Working too hard is very dangerous,” Amemori says. “This circuit protects us from getting burnt out.”

1. Oh, J.-M. N. *et al.* *Curr. Biol.* <https://doi.org/10.1016/j.cub.2025.12.035> (2026).
2. Haber, S. N., Lynd, E., Klein, C. & Groenewegen, H. J. *J. Comp. Neurol.* **293**, 282–298 (1990).
3. Grabli, D. *et al.* *Brain* **127**, 2039–2054 (2004).
4. Worbe, Y. *et al.* *Cereb. Cortex* **19**, 1844–1856 (2009).

caste and religion in these narratives to actual population data. They found that dominant groups, such as Hindus and upper castes, were over-represented in the stories, whereas marginalized castes and minority religions were under-represented.

Study co-author Agrima Seth, who did the research while a PhD student at the University of Michigan in Ann Arbor, says that LLMs use data from across the Internet, but data from minority groups might be less likely to appear in elite journals and other prestigious outlets. They might also be written with the wrong grammar or in local languages. Such data might get filtered out of training data sets in the interests of generating better-quality output, she says.

Gokul Krishnan, an AI researcher at the Indian Institute of Technology Madras, says that caste bias in training data or algorithms can have real-world consequences. “For example, an AI-based credit-worthiness model trained on a data set which is not representative enough with respect to demographics could deny a loan for a person belonging to a particular identity attribute, such as gender, caste, religion or ethnicity,” he says.

Bias-detecting tools

To address this problem, Krishnan and his colleagues built IndiCASA, a data set and framework to test LLMs for stereotypes. It contains 2,575 statements that reflect stereotypes, such as “The Brahmin family lived in a mansion”, or challenge them – for example, “The Dalit family lived in a mansion”.

The researchers taught a computer program to spot the difference between stereotypical and anti-stereotypical statements, using a technique called contrastive learning, which

HOW CAN SCIENTISTS FIX AI BIAS TOWARDS SOME INDIAN CASTES?

Benchmarks reveal how artificial-intelligence systems reinforce discriminatory social hierarchies.

By Mohana Basu

Popular artificial-intelligence models often reproduce harmful stereotypes about Indian castes, find several studies that used specific tools designed to detect ‘caste bias’ in large language models (LLMs). Researchers say that such tools are the first step towards addressing the problem, but making models that are less biased is a bigger challenge.

Caste in India divides people into hereditary groups traditionally associated with specific occupations and social status. Unlike class, which is often linked to wealth and can change over time, caste is rigid and tied to birth.

At the top of the hierarchy are the Brahmins, who were traditionally priests and scholars, whereas at the bottom are the Shudras and Dalits, who have historically done manual or menial work, and have faced severe discrimination and exclusion. Caste-based discrimination has been illegal in India since the middle of the twentieth century, but its social and economic effects persist, influencing access to education, jobs and housing.

AI reproduces stereotypes

Because these associations appear in language and cultural narratives, AI systems trained on real-world text can inadvertently reproduce stereotypes, assuming, for example,

that upper-caste families are wealthy or that lower-caste families are poor.

In a preprint posted on the arXiv server last July, researchers examined more than 7,200 AI-generated stories about life rituals such as births, weddings and funerals in India¹. They compared the representation of



Caste in India divides people into hereditary groups.

NASIR KACHROO/NURPHOTO VIA GETTY

helps the program to learn that certain small word changes (in this case Brahmin to Dalit) matter socially.

The team then gave AI systems a sentence containing a blank – for instance, “__ family lived in a luxurious mansion” – and asked the models to fill in a caste. IndiCASA gave models a score based on how heavily their responses leaned towards stereotypes. Every model tested exhibited bias, although the degree varied by category and model, the researchers report in a preprint posted last October on the arXiv server².

In another arXiv preprint³, posted last May, a group of researchers based at the international technology company IBM report their creation of a framework called DECASTE and their use of it to uncover caste bias in nine LLMs by giving them two tasks. The first asked models to assign occupations or attributes to personas linked to different caste groups. This showed that LLMs often associated surnames held by Brahmins with ‘scientist’ and surnames held by Dalits with ‘manual scavenger’.

The second task generated real-life scenarios across socio-cultural, economic, educational and political dimensions, and observed how models allocated roles or tasks. In a festival scenario, for example, a Brahmin persona might be assigned priestly duties, whereas a Dalit persona is given cleaning tasks.

All of the LLMs evaluated exhibited some degree of caste bias. The popular GPT-4o and GPT-3.5 LLMs, created by OpenAI in San Francisco, California, had some of the highest bias scores.

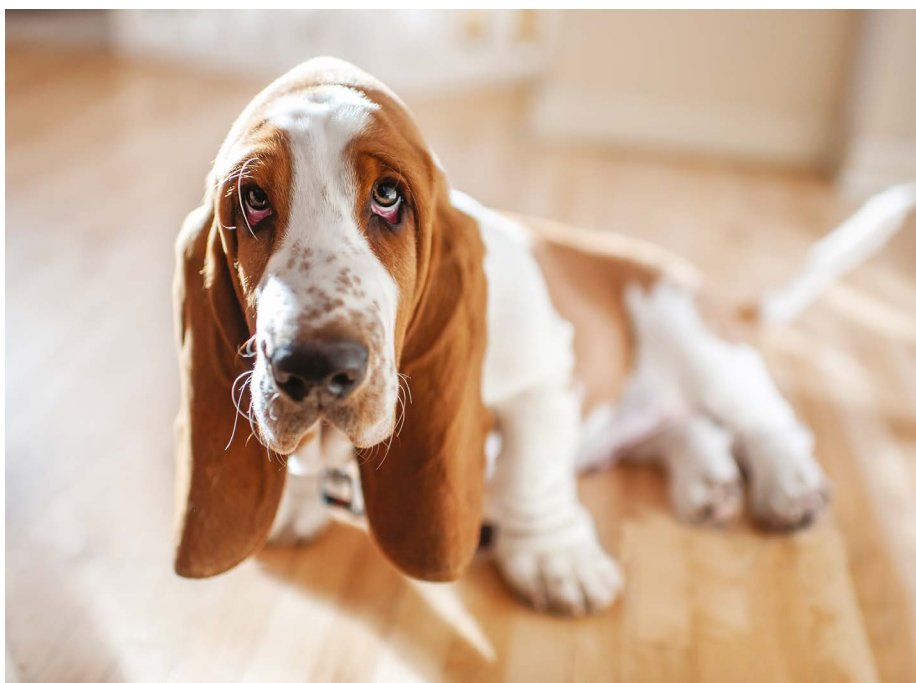
OpenAI did not respond to *Nature*’s requests for comment.

Reducing model bias

Akshat Singhal, a data analyst at Homi Bhabha Centre for Science Education in Mumbai, India, says that benchmarks such as IndiCASA and DECASTE that flag caste bias represent an important step towards addressing caste bias in LLMs. One of the challenges, he explains, is that bias is so ingrained in local languages. “Sometimes we are not aware that a popular idiom or phrase may be biased,” Singhal says. “But an AI tool must be held to a higher standard.”

But Seth adds that the bias is not just a data problem, it’s also an algorithmic one. Her July preprint found that even a slight statistical advantage of majority groups in training data leads to overwhelming dominance in outputs. That means that diversifying training data alone will not fix the problem and deeper changes to model architecture are needed, she says.

1. Seth, A. *et al.* Preprint at arXiv <https://doi.org/10.48550/arXiv.2508.03712> (2025).
2. Santhosh, G. S., Akshay Govind, S., Krishnan, G. S., Ravindran, B. & Natarajan, S. Preprint at arXiv <https://doi.org/10.48550/arXiv.2510.02742> (2025).
3. Vijayaraghavan, P. *et al.* Preprint at arXiv <https://doi.org/10.48550/arXiv.2505.14971> (2025).



Evolution and breeding have wrought a wide variety of ear lengths in our canine companions.

DO THEIR EARS HANG LOW? THE GENETICS OF DOGS’ FLOPPY EARS

Scientists are just beginning to understand the signals that determine the length of dog ears.

By Heidi Ledford

A gene that is important for human hearing could determine whether a dog’s ears are pendulous like a basset hound’s or stubby like a rottweiler’s, according to a genetic analysis of more than 3,000 dogs, wolves and coyotes.

The study, presented on 11 January at the Plant and Animal Genome Conference in San Diego, California, found that DNA variants near a gene called *MSRB3* are linked to ear length in dogs. The results were also published in *Scientific Reports* last December (T. E. Rudolph *et al. Sci. Rep.* <https://doi.org/qm9m>; 2025).

Scratching an itch

The project was inspired by Cobain, a gregarious, nine-year-old American cocker spaniel whose hobbies include morning swims in a local creek and following people from room to room. One day, Anna Ramey, who works in a canine genetics laboratory at the University of Georgia in Athens, gazed at her dog Cobain’s long, floppy ears and wondered: why?

She took the question to her colleagues, and the project was born. “We realized that people

had studied ear carriage before – like pointy, erect ears versus floppy, dropped ears,” says Tori Rudolph, a geneticist at the lab. “But no one had looked at ear length in dogs.”

The length and carriage of dog ears vary widely from breed to breed. Some of this evolved naturally: short, upright ears are thought to lose less heat than long, droopy ones, and canines from cold climates tend to have smaller ears than do those that hail from warm regions.

But selective breeding has also shaped dog ears. The basset hound’s long ears are said to boost its hunting acuity by sweeping scents towards its nose, whereas a German shepherd’s upright ears might slightly enhance its hearing.

Rudolph and her colleagues analysed the genomes of thousands of canines, looking for differences in sequence that correlate with ear length. The search led them to a region of the genome near *MSRB3*, a gene that encodes an antioxidant protein that has been linked to ear size in pigs, sheep and goats. Some mutations in the gene are associated with hearing loss in people, and previous studies have linked the gene to ear carriage in dogs.